

Miniversal deformations of matrices of sesquilinear forms

Andrii R. Dmytryshyn* Vyacheslav Futorny[†]
 Vladimir V. Sergeichuk[‡]

Abstract

V.I. Arnold [Russian Math. Surveys 26 (2) (1971) 29–43] constructed a miniversal deformation of matrices under similarity; that is, a simple normal form to which not only a given square matrix A but all matrices B close to it can be reduced by similarity transformations that smoothly depend on the entries of B . A miniversal deformation of matrices under congruence was constructed by V. Futorny and V.V. Sergeichuk [Miniversal deformations of matrices of bilinear forms, Preprint RT-MAT 2007-04, Universidade de São Paulo, 2007, 34 p. (arXiv:1004.3584v1)]. We similarly construct miniversal deformation of matrices under *congruence.

AMS classification: 15A21

Keywords: Sesquilinear form; Classification; Miniversal deformation

*Faculty of Mechanics and Mathematics, Kiev National Taras Shevchenko University, Volodymyrska 64, Kiev, Ukraine. Email: AndriiDmytryshyn@gmail.com.

[†]Department of Mathematics, University of São Paulo, Brazil. Email: futorny@ime.usp.br. Supported in part by the CNPq grant (301743/2007-0) and by the Fapesp grant (2010/50347-9).

[‡]Institute of Mathematics, Tereshchenkivska 3, Kiev, Ukraine. Email: sergeich@imath.kiev.ua. Supported in part by the Fapesp grant (2010/07278-6). The work was done while this author was visiting the University of São Paulo, whose hospitality is gratefully acknowledged.

1 Introduction

The reduction of a matrix to its Jordan form is an unstable operation: both the Jordan form and the reduction transformations depend discontinuously on the entries of the original matrix. Therefore, if the entries of a matrix are known only approximately, then it is unwise to reduce it to Jordan form. Furthermore, when investigating a family of matrices smoothly depending on parameters, then although each individual matrix can be reduced to a Jordan form, it is unwise to do so since in such an operation the smoothness relative to the parameters is lost.

For these reasons V.I. Arnold [1] (see also [2, 3]) constructed miniversal deformations of matrices under similarity; that is, a simple normal form to which not only a given square matrix A but all matrices B close to it can be reduced by similarity transformations that smoothly depend on the entries of B . Miniversal deformations were also constructed for:

- (a) real matrices with respect to similarity by Galin [11] (see also [2, 3]); his normal form was simplified in [13];
- (b) complex matrix pencils by Edelman, Elmroth, and Kågström [8]; a simpler normal form of complex and real matrix pencils was constructed in [13];
- (c) complex and real contragredient matrix pencils (i.e., matrices of pairs of counter linear operators $U \rightleftarrows V$) in [13];
- (d) matrices of linear operators on a unitary space by Benedetti and Crag-nolini [4]; and
- (e) matrices of selfadjoint operators on a complex or real vector space with scalar product given by a skew-symmetric, or symmetric, or Hermitian nonsingular form in [12, 7, 21, 23].

Futorny and Sergeichuk [9] constructed a miniversal deformation of matrices of complex bilinear forms; that is, of matrices under *congruence transformations*

$$A \mapsto S^T A S, \quad S \text{ is nonsingular}$$

(and also miniversal deformations of pairs consisting of symmetric and skew-symmetric matrices since each square matrix is their sum).

In this paper, we construct an analogous miniversal deformation of matrices of complex sesquilinear forms; that is, of matrices under **congruence transformations*

$$A \mapsto S^*AS, \quad S \text{ is nonsingular}$$

(and also miniversal deformations of pairs $(\mathcal{H}, \mathcal{K})$ of Hermitian matrices since each square matrix is uniquely represented in the form $\mathcal{H} + i\mathcal{K}$; see Remark 3.1).

All matrices that we consider are complex matrices. In Sections 2 and 3, we give miniversal deformations of matrices of bilinear forms. In Sections 4–7 we prove that these deformations are miniversal.

2 The main theorem in terms of holomorphic functions

Define the n -by- n matrices:

$$J_n(\lambda) := \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}, \quad \Delta_n = \begin{bmatrix} 0 & & & 1 \\ & \ddots & & i \\ & & 1 & \ddots \\ 1 & i & & 0 \end{bmatrix}.$$

The most important property of the symmetric matrices Δ_n is that $\Delta_n^{-*} \Delta_n = \bar{\Delta}_n^{-1} \Delta_n$ is similar to $J_n(1)$.

We use the following canonical form of complex matrices for **congruence*.

Theorem 2.1. *Each square complex matrix is *congruent to a direct sum, uniquely determined up to permutation of summands, of canonical matrices of the three types*

$$H_m(\lambda) := \begin{bmatrix} 0 & I_m \\ J_m(\lambda) & 0 \end{bmatrix} \quad (|\lambda| > 1), \quad \mu \Delta_n \quad (|\mu| = 1), \quad J_k(0) \quad (1)$$

in which $\lambda, \mu \in \mathbb{C}$.

This canonical form was obtained in [15] basing on [24, Theorem 3] and generalized to other fields in [18]; a direct proof that this form is canonical is given in [16, 17].

Let A be a given n -by- n matrix, and let

$$A_{\text{can}} = \bigoplus_i H_{p_i}(\lambda_i) \oplus \bigoplus_j \mu_j \Delta_{q_j} \oplus \bigoplus_l J_{r_l}(0), \quad r_1 \geq r_2 \geq \dots, \quad (2)$$

be its canonical form for $*$ -congruence. All matrices that are close to A are represented in the form $A + E$ in which $E \in \mathbb{C}^{n \times n}$ is close to 0_n . Let $\mathcal{S}(E)$ be a holomorphic $n \times n$ matrix function in some neighborhood of 0 (this means that each of its entries is a power series in the entries of E that is convergent in this neighborhood of 0). Define $\mathcal{D}(E)$ by

$$A_{\text{can}} + \mathcal{D}(E) = \mathcal{S}(E)^*(A + E)\mathcal{S}(E), \quad \mathcal{S}(0) = S. \quad (3)$$

Then $\mathcal{D}(E)$ is holomorphic at 0 and $\mathcal{D}(0) = 0$. In the next theorem we obtain $\mathcal{D}(E)$ with the minimal number of nonzero entries that can be attained by using transformations (3). By a $(0, *, \circ, \bullet)$ *matrix* we mean a matrix whose entries are of the form 0, $*$, \circ , and \bullet . The theorem involves the following $(0, *, \circ, \bullet)$ matrices:

- The $m \times n$ matrices

$$0^{\swarrow} := \begin{bmatrix} * & & \\ \vdots & & \\ * & & \end{bmatrix} \quad 0 \quad \text{if } m \leq n \text{ or } \begin{bmatrix} & & \\ & 0 & \\ * & \dots & * \end{bmatrix} \quad \text{if } m \geq n,$$

$$0^{\wedge} := \begin{bmatrix} \vdots & & \\ 0 & & \\ * & & \\ 0 & & \\ * & & \end{bmatrix} \quad 0 \quad \text{if } m \leq n \text{ or } \begin{bmatrix} & & & & \\ & & 0 & & \\ & & & & \\ & & & & \\ * & 0 & * & 0 & \dots \end{bmatrix} \quad \text{if } m \geq n,$$

(choosing among the left and right matrices in these equalities, we take a matrix with the minimal number of stars; we can take any of them if $m = n$).

- The matrices

$$0^{\nwarrow}, \quad 0^{\nearrow}, \quad 0^{\searrow}$$

that are obtained from 0^{\swarrow} by the clockwise rotation through 90° , 180° , and 270° .

- The $n \times n$ matrices

$$0^{\searrow} := \begin{cases} \text{diag}(*, \dots, *, 0, \dots, 0) & \text{if } n = 2k, \\ \text{diag}(*, \dots, *, \circ, 0, \dots, 0) & \text{if } n = 2k + 1, \end{cases} \quad (4)$$

$$0^{\swarrow} := \begin{cases} \text{diag}(*, \dots, *, 0, \dots, 0) & \text{if } n = 2k, \\ \text{diag}(*, \dots, *, \bullet, 0, \dots, 0) & \text{if } n = 2k + 1. \end{cases} \quad (5)$$

in which the number of $*$'s is equal to k .

- The $m \times n$ matrices

$$0^\dagger := \begin{bmatrix} * & \cdots & * \\ & & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} & & 0 \\ * & \cdots & * \end{bmatrix}$$

(0^\dagger can be taken in any of these forms), and

$$\mathcal{P}_{mn} := \begin{bmatrix} 0 & \cdots & 0 & & 0 \\ \vdots & \ddots & \vdots & & \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix} \quad \begin{array}{l} \text{with } m \leq n \\ \text{and } n - m - 1 \text{ stars.} \end{array} \quad (6)$$

Let $A_{\text{can}} = A_1 \oplus A_2 \oplus \cdots \oplus A_t$ be the decomposition (2). Partition \mathcal{D} in (3) conformably to the partition of A_{can} :

$$\mathcal{D} = \mathcal{D}(E) = \begin{bmatrix} \mathcal{D}_{11} & \cdots & \mathcal{D}_{1t} \\ \vdots & \ddots & \vdots \\ \mathcal{D}_{t1} & \cdots & \mathcal{D}_{tt} \end{bmatrix}, \quad (7)$$

and write

$$\mathcal{D}(A_i) := \mathcal{D}_{ii}, \quad \mathcal{D}(A_i, A_j) := (\mathcal{D}_{ji}, \mathcal{D}_{ij}) \quad \text{if } i < j. \quad (8)$$

Our main result is the following theorem, which we reformulate in a more abstract form in Theorem 3.1.

Theorem 2.2. *Let A be a square complex matrix and let (2) be its canonical matrix for congruence. All matrices $A + E$ that are sufficiently close to A can be simultaneously reduced by transformations*

$$A + E \mapsto \mathcal{S}(E)^*(A + E)\mathcal{S}(E), \quad \begin{array}{l} \mathcal{S}(E) \text{ is nonsingular} \\ \text{and holomorphic at zero,} \end{array} \quad (9)$$

to the form $A_{can} + \mathcal{D}$ in which \mathcal{D} is a $(0, *, \circ, \bullet)$ matrix such that the number of zero entries in \mathcal{D} is maximal that can be achieved by transformations (9), the symbols $*$, \circ , and \bullet in \mathcal{D} represent complex, real, and pure imaginary entries that depend holomorphically on the entries of E , and the blocks of \mathcal{D} with respect to the partition (7) are defined in the notation (8) by the following equalities in which $|\lambda| > 1$, $|\lambda'| > 1$, and $|\mu| = |\mu'| = 1$:

(i) The diagonal blocks of \mathcal{D} are defined by

$$\mathcal{D}(H_m(\lambda)) = \begin{bmatrix} 0 & 0 \\ 0^\swarrow & 0 \end{bmatrix}; \quad (10)$$

$$\mathcal{D}(\mu\Delta_n) = \begin{cases} 0^\circ & \text{if } \mu \notin \mathbb{R}, \\ 0^\searrow & \text{if } \mu \notin i\mathbb{R} \end{cases} \quad (11)$$

(if $\mu \notin \mathbb{R} \cup i\mathbb{R}$ then we can use both 0° and 0^\searrow);

$$\mathcal{D}(J_n(0)) = 0^\wedge. \quad (12)$$

(ii) The off-diagonal blocks of \mathcal{D} whose horizontal and vertical strips contain summands of A_{can} of the same type are defined by

$$\mathcal{D}(H_m(\lambda), H_n(\lambda')) = \begin{cases} (0, 0) & \text{if } \lambda \neq \lambda', \\ \left(\begin{bmatrix} 0 & 0^\nearrow \\ 0^\swarrow & 0 \end{bmatrix}, 0 \right) & \text{if } \lambda = \lambda'; \end{cases} \quad (13)$$

$$\mathcal{D}(\mu\Delta_m, \mu'\Delta_n) = \begin{cases} (0, 0) & \text{if } \mu \neq \pm\mu', \\ (0^\wedge, 0) & \text{if } \mu = \pm\mu'; \end{cases} \quad (14)$$

$$\mathcal{D}(J_m(0), J_n(0)) = \begin{cases} (0^\wedge, 0^\wedge) & \text{if } m \geq n \text{ and } n \text{ is even,} \\ (0^\wedge + \mathcal{P}_{nm}, 0^\wedge) & \text{if } m \geq n \text{ and } n \text{ is odd.} \end{cases} \quad (15)$$

(iii) The off-diagonal blocks of \mathcal{D} whose horizontal and vertical strips contain summands of A_{can} of different types are defined by

$$\mathcal{D}(H_m(\lambda), \mu\Delta_n) = (0, 0); \quad (16)$$

$$\mathcal{D}(H_m(\lambda), J_n(0)) = \mathcal{D}(\mu\Delta_m, J_n(0)) = \begin{cases} (0, 0) & \text{if } n \text{ is even,} \\ (0^\dagger, 0) & \text{if } n \text{ is odd.} \end{cases} \quad (17)$$

For each $A \in \mathbb{C}^{n \times n}$, the set

$$T(A) := \{C^* A + AC \mid C \in \mathbb{C}^{n \times n}\} \quad (18)$$

is a vector space over \mathbb{R} , which is the tangent space to the congruence class of A at the point A since

$$(I + \varepsilon C)^* A (I + \varepsilon C) = A + \varepsilon(C^* A + AC) + \varepsilon^2 C^* AC \quad (19)$$

for all $C \in \mathbb{C}^{n \times n}$ and $\varepsilon \in \mathbb{R}$.

The matrix \mathcal{D} from Theorem 2.2 was constructed such that

$$\mathbb{C}^{n \times n} = T(A_{\text{can}}) \oplus_{\mathbb{R}} \mathcal{D}(\mathbb{C}) \quad (20)$$

in which $\mathcal{D}(\mathbb{C})$ is the vector space of all matrices obtained from \mathcal{D} by replacing its entries $*$, \circ , and \bullet in \mathcal{D} by complex, real, and pure imaginary numbers. Thus, the double number of stars plus the number of circles plus the number of bullets in \mathcal{D} is the codimension over \mathbb{R} of the $*$ -congruence class of A_{can} ; it was independently calculated in [6]. Simplest miniversal deformations of matrix pencils and contagredient matrix pencils and of matrices under congruence were constructed in [13, 9] by an analogous method.

Theorem 2.2 will be proved in Sections 4–7 as follows: we first prove in Lemma 4.2 that each $(0, *, \circ, \bullet)$ matrix that satisfies (20) can be taken as \mathcal{D} in Theorem 2.2, and then verify that \mathcal{D} from Theorem 2.2 satisfies (20).

3 The main theorem in terms of miniversal deformations

The notion of a miniversal deformation of a matrix with respect similarity was given by V. I. Arnold [1] (see also [3, §30B]) and can be extended to matrices with respect to $*$ -congruence as follows.

A *deformation* of a matrix $A \in \mathbb{C}^{n \times n}$ is a holomorphic map $\mathcal{A}: \Lambda \rightarrow \mathbb{C}^{n \times n}$ in which $\Lambda \subset \mathbb{R}^k$ is a neighborhood of $\vec{0} = (0, \dots, 0)$ and $\mathcal{A}(\vec{0}) = A$.

Let \mathcal{A} and \mathcal{B} be two deformations of A with the same parameter space \mathbb{R}^k . \mathcal{A} and \mathcal{B} are considered as *equal* if they coincide on some neighborhood of $\vec{0}$ (this means that each deformation is a germ). We say that \mathcal{A} and \mathcal{B} are *equivalent* if the identity matrix I_n possesses a deformation \mathcal{I} such that

$$\mathcal{B}(\vec{\lambda}) = \mathcal{I}(\vec{\lambda})^* \mathcal{A}(\vec{\lambda}) \mathcal{I}(\vec{\lambda}) \quad (21)$$

for all $\vec{\lambda} = (\lambda_1, \dots, \lambda_k)$ in some neighborhood of $\vec{0}$.

Definition 3.1. A deformation $\mathcal{A}(\lambda_1, \dots, \lambda_k)$ of a matrix A is called *versal* if every deformation $\mathcal{B}(\mu_1, \dots, \mu_l)$ of A is equivalent to a deformation of the form $\mathcal{A}(\varphi_1(\vec{\mu}), \dots, \varphi_k(\vec{\mu}))$, in which all $\varphi_i : \mathbb{R} \rightarrow \mathbb{R}$ are power series that are convergent in a neighborhood of $\vec{0}$ and $\varphi_i(\vec{0}) = 0$. A versal deformation $\mathcal{A}(\lambda_1, \dots, \lambda_k)$ of A is called *miniversal* if there is no versal deformation having less than k parameters.

For each $(0, *, \circ, \bullet)$ matrix \mathcal{D} , we denote by $\mathcal{D}(\mathbb{C})$ the real space of all matrices obtained from \mathcal{D} by replacing its entries $*$, \circ , and \bullet by complex, real, and pure imaginary numbers (as in (20)) and by $\mathcal{D}(\vec{\varepsilon})$ the parameter matrix obtained from \mathcal{D} by replacing each (i, j) star with $\varepsilon_{ij} + i\varepsilon'_{ij}$, each (i, j) circle with ε_{ij} , and each (i, j) bullet with $i\varepsilon'_{ij}$. This means that

$$\mathcal{D}(\mathbb{C}) := \left(\bigoplus_{(i,j) \in \mathcal{I}_*(\mathcal{D})} \mathbb{C}E_{ij} \right) \oplus \left(\bigoplus_{(i,j) \in \mathcal{I}_\circ(\mathcal{D})} \mathbb{R}E_{ij} \right) \oplus \left(\bigoplus_{(i,j) \in \mathcal{I}_\bullet(\mathcal{D})} i\mathbb{R}E_{ij} \right), \quad (22)$$

$$\mathcal{D}(\vec{\varepsilon}) := \left(\sum_{(i,j) \in \mathcal{I}_*(\mathcal{D})} (\varepsilon_{ij} + i\varepsilon'_{ij})E_{ij} \right) + \left(\sum_{(i,j) \in \mathcal{I}_\circ(\mathcal{D})} \varepsilon_{ij}E_{ij} \right) + \left(\sum_{(i,j) \in \mathcal{I}_\bullet(\mathcal{D})} i\varepsilon'_{ij}E_{ij} \right), \quad (23)$$

where

$$\mathcal{I}_*(\mathcal{D}), \mathcal{I}_\circ(\mathcal{D}), \mathcal{I}_\bullet(\mathcal{D}) \subseteq \{1, \dots, n\} \times \{1, \dots, n\} \quad (24)$$

are the sets of indices of the stars of the circles, and of the bullets in \mathcal{D} , and E_{ij} is the elementary matrix whose (i, j) entry is 1 and the others are 0.

We say that a miniversal deformation of A is *simplest* if it has the form $A + \mathcal{D}(\vec{\varepsilon})$, where \mathcal{D} is a $(0, *, \circ, \bullet)$ matrix. If all entries of \mathcal{D} are stars, then it defines the deformation

$$\mathcal{U}(\vec{\varepsilon}) := A + \sum_{i,j=1}^n (\varepsilon_{ij} + i\varepsilon'_{ij})E_{ij}. \quad (25)$$

Since each square matrix is $*$ -congruent to its canonical matrix, it suffices to construct miniversal deformations of canonical matrices (2). These deformations are given in the following theorem, which is a stronger form of Theorem 2.2.

Theorem 3.1. *Let A_{can} be a canonical matrix (2) for congruence. A simplest miniversal deformation of A_{can} can be taken in the form $A_{can} + \mathcal{D}(\vec{\varepsilon})$, where \mathcal{D} is the $(0, *, \circ, \bullet)$ matrix partitioned into blocks \mathcal{D}_{ij} (as in (7)) that are defined by (10)–(17) in the notation (8).*

Remark 3.1. Theorem 3.1 also gives a miniversal deformation of a canonical pair for \ast -congruence of Hermitian matrices $(\mathcal{H}_{\text{can}}, \mathcal{K}_{\text{can}})$ of the same size; that is, a normal form with minimal number of parameters to which all pairs of Hermitian matrices $(\mathcal{H}, \mathcal{K})$ that are close to $(\mathcal{H}_{\text{can}}, \mathcal{K}_{\text{can}})$ can be reduced by \ast -congruence transformations

$$(\mathcal{H}, \mathcal{K}) \mapsto (S^* \mathcal{H} S, S^* \mathcal{K} S), \quad S \text{ is nonsingular,}$$

in which S smoothly depends on the entries of \mathcal{H} and \mathcal{K} . All one has to do is to express $A_{\text{can}} + \mathcal{D}(\vec{\varepsilon})$ as the sum $\mathcal{H}(\vec{\varepsilon}) + i\mathcal{K}(\vec{\varepsilon})$ in which $\mathcal{H}(\vec{\varepsilon})$ and $\mathcal{K}(\vec{\varepsilon})$ are Hermitian matrices. The canonical pair $(\mathcal{H}_{\text{can}}, \mathcal{K}_{\text{can}})$ such that $\mathcal{H}_{\text{can}} + i\mathcal{K}_{\text{can}} = A_{\text{can}}$ was described in [17, Theorem 1.2(b)].

4 Beginning of the proof of Theorem 3.1

Let us give a method of constructing simplest miniversal deformations, which is used in the proof of Theorem 3.1.

The deformation (25) is universal in the sense that every deformation $\mathcal{B}(\mu_1, \dots, \mu_l)$ of A has the form $\mathcal{U}(\vec{\varphi}(\mu_1, \dots, \mu_l))$, where $\varphi_{ij} : \mathbb{R} \rightarrow \mathbb{R}$ are power series that are convergent in a neighborhood of $\vec{0}$ and $\varphi_{ij}(\vec{0}) = 0$. Hence every deformation $\mathcal{B}(\mu_1, \dots, \mu_l)$ in Definition 3.1 can be replaced by $\mathcal{U}(\vec{\varepsilon})$, which gives the following lemma.

Lemma 4.1. *The following two conditions are equivalent for any deformation $\mathcal{A}(\lambda_1, \dots, \lambda_k)$ of a matrix A :*

- (i) *The deformation $\mathcal{A}(\lambda_1, \dots, \lambda_k)$ is versal.*
- (ii) *The deformation (25) is equivalent to $\mathcal{A}(\varphi_1(\vec{\varepsilon}), \dots, \varphi_k(\vec{\varepsilon}))$ for some power series $\varphi_i : \mathbb{R} \rightarrow \mathbb{R}$ that are convergent in a neighborhood of $\vec{0}$ and such that $\varphi_i(\vec{0}) = 0$.*

If U is a subspace of a vector space V , then each set $v + U$ with $v \in V$ is called an *affine subspace parallel to U* .

The proof of Theorem 3.1 is based on the following lemma, which gives a method of constructing miniversal deformations. A constructive proof of this lemma is given in Section ??.

Lemma 4.2. *Let $A \in \mathbb{C}^{n \times n}$ and let \mathcal{D} be a $(0, \ast, \circ, \bullet)$ matrix of size $n \times n$. The following three statements are equivalent:*

(i) The deformation $A + \mathcal{D}(\varepsilon)$ defined in (23) is miniversal.

(ii) The vector space $\mathbb{C}^{n \times n}$ decomposes into the direct sum

$$\mathbb{C}^{n \times n} = T(A) \oplus_{\mathbb{R}} \mathcal{D}(\mathbb{C}) \quad (26)$$

in which $T(A)$ is the vector space over \mathbb{R} defined in (18).

(iii) Each affine \mathbb{R} -subspace in $\mathbb{C}^{n \times n}$ parallel to $T(A)$ intersects $\mathcal{D}(\mathbb{C})$ at exactly one point.

Proof. Define the action of the group $GL_n(\mathbb{C})$ of nonsingular n -by- n matrices on the space $\mathbb{C}^{n \times n}$ by

$$A^S := S^* A S, \quad A \in \mathbb{C}^{n \times n}, \quad S \in GL_n(\mathbb{C}). \quad (27)$$

The orbit A^{GL_n} of A under this action consists of all matrices that are *-congruent to A .

By (19), the space $T(A)$ is the tangent space to the orbit A^{GL_n} at the point A . Hence $\mathcal{D}(\varepsilon)$ is transversal to the orbit A^{GL_n} at the point A if

$$\mathbb{C}^{n \times n} = T(A) + \mathcal{D}(\mathbb{C})$$

(see definitions in [3, §29E]; two subspaces of a vector space are called *transversal* if their sum is equal to the whole space).

This proves the equivalence of (i) and (ii) since a transversal (of the minimal dimension) to the orbit is a (mini)versal deformation [2, Section 1.6]. The equivalence of (ii) and (iii) is obvious. \square

Recall that the orbits of canonical matrices (2) under the action (27) were also studied in [6].

Due to Lemma 4.2, a simplest miniversal deformation of $A \in \mathbb{C}^{n \times n}$ can be constructed as follows. Let T_1, \dots, T_r be a basis of the space $T(A)$, and let $E_1, \dots, E_{n^2}, iE_1, \dots, iE_{n^2}$ be the basis of $\mathbb{C}^{n \times n}$ over \mathbb{R} , in which E_1, \dots, E_{n^2} are all elementary matrices E_{ij} . Removing from the sequence $T_1, \dots, T_r, E_1, \dots, E_{n^2}, iE_1, \dots, iE_{n^2}$ every matrix that is a linear combination of the preceding matrices, we obtain a new basis $T_1, \dots, T_r, E_{i_1}, \dots, E_{i_k}, E_{j_1}, \dots, E_{j_\ell}$ of the space $\mathbb{C}^{n \times n}$ over \mathbb{R} . By Lemma 4.2, the deformation

$$\mathcal{A}(\varepsilon_1, \dots, \varepsilon_k, \varepsilon'_1, \dots, \varepsilon'_\ell) = A + \varepsilon_1 E_{i_1} + \dots + \varepsilon_k E_{i_k} + \varepsilon'_1 E_{j_1} + \dots + \varepsilon'_\ell E_{j_\ell}$$

is miniversal.

For each $M \in \mathbb{C}^{m \times m}$ and $N \in \mathbb{C}^{n \times n}$, define the real vector space

$$T(M, N) := \{(S^*M + NR, R^*N + MS) \mid S \in \mathbb{C}^{m \times n}, R \in \mathbb{C}^{n \times m}\}. \quad (28)$$

Lemma 4.3. *Let $A = A_1 \oplus \dots \oplus A_t$ be a block-diagonal matrix in which every A_i is $n_i \times n_i$. Let $\mathcal{D} = [\mathcal{D}_{ij}]$ be a $(0, *, \circ, \bullet)$ matrix of the same size and partitioned into blocks conformably to the partition of A . Then $A + \mathcal{D}(\varepsilon)$ is a simplest miniversal deformation of A for congruence if and only if*

- (i) *each affine \mathbb{R} -subspace in $\mathbb{C}^{n_i \times n_i}$ parallel to $T(A_i)$ (defined in (18)) intersects $\mathcal{D}_{ii}(\mathbb{C})$ at exactly one point, and*
- (ii) *each affine \mathbb{R} -subspace in $\mathbb{C}^{n_j \times n_i} \oplus \mathbb{C}^{n_i \times n_j}$ parallel to $T(A_i, A_j)$ intersects $\mathcal{D}_{ji}(\mathbb{C}) \oplus \mathcal{D}_{ij}(\mathbb{C})$ at exactly one point.*

Proof. By Lemma 4.2(iii), $A + \mathcal{D}(\varepsilon)$ is a simplest miniversal deformation of A if and only if for each $C \in \mathbb{C}^{n \times n}$ the affine \mathbb{R} -subspace $C + T(A)$ contains exactly one $D \in \mathcal{D}(\mathbb{C})$; that is, exactly one

$$D = C + S^*A + AS \in \mathcal{D}(\mathbb{C}) \quad \text{with } S \in \mathbb{C}^{n \times n}. \quad (29)$$

Partition D , C , and S into blocks conformably to the partition of A . By (29), for each i we have $D_{ii} = C_{ii} + S_{ii}^*A_{ii} + A_{ii}S_{ii}$, and for all i and j such that $i < j$ we have

$$\begin{bmatrix} D_{ii} & D_{ij} \\ D_{ji} & D_{jj} \end{bmatrix} = \begin{bmatrix} C_{ii} & C_{ij} \\ C_{ji} & C_{jj} \end{bmatrix} + \begin{bmatrix} S_{ii}^* & S_{ji}^* \\ S_{ij}^* & S_{jj}^* \end{bmatrix} \begin{bmatrix} A_i & 0 \\ 0 & A_j \end{bmatrix} + \begin{bmatrix} A_i & 0 \\ 0 & A_j \end{bmatrix} \begin{bmatrix} S_{ii} & S_{ij} \\ S_{ji} & S_{jj} \end{bmatrix}.$$

Thus, (29) is equivalent to the conditions

$$D_{ii} = C_{ii} + S_{ii}^*A_i + A_iS_{ii} \in \mathcal{D}_{ii}(\mathbb{C}) \quad \text{for } 1 \leq i \leq t \quad (30)$$

and

$$(D_{ji}, D_{ij}) = (C_{ji}, C_{ij}) + (S_{ij}^*A_i + A_jS_{ji}, S_{ji}^*A_j + A_iS_{ij}) \in \mathcal{D}_{ji}(\mathbb{C}) \oplus \mathcal{D}_{ij}(\mathbb{C}) \quad (31)$$

for $1 \leq i < j \leq t$. Hence for each $C \in \mathbb{C}^{n \times n}$ there exists exactly one $D \in \mathcal{D}$ of the form (29) if and only if

- (i') for each $C_{ii} \in \mathbb{C}^{n_i \times n_i}$ there exists exactly one $D_{ii} \in \mathcal{D}_{ii}$ of the form (30), and

(ii') for each $(C_{ji}, C_{ij}) \in \mathbb{C}^{n_j \times n_i} \oplus \mathbb{C}^{n_i \times n_j}$ there exists exactly one $(D_{ji}, D_{ij}) \in \mathcal{D}_{ji}(\mathbb{C}) \oplus \mathcal{D}_{ij}(\mathbb{C})$ of the form (31).

This proves the lemma. \square

Corollary 4.1. In the notation of Lemma 4.3, $A + \mathcal{D}(\vec{\varepsilon})$ is a miniversal deformation of A if and only if each submatrix of the form

$$\begin{bmatrix} A_i + \mathcal{D}_{ii}(\vec{\varepsilon}) & \mathcal{D}_{ij}(\vec{\varepsilon}) \\ \mathcal{D}_{ji}(\vec{\varepsilon}) & A_j + \mathcal{D}_{jj}(\vec{\varepsilon}) \end{bmatrix} \quad \text{with } i < j$$

is a miniversal deformation of $A_i \oplus A_j$. A similar reduction to the case of canonical forms for congruence with two direct summands was used in [6] for the solution of the equation $XA + AX^* = 0$.

Let us start to prove Theorem 2.2. Let $A_{\text{can}} = A_1 \oplus A_2 \oplus \cdots \oplus A_t$ be the canonical matrix (2), and let $\mathcal{D} = [\mathcal{D}_{ij}]_{i,j=1}^t$ be the $(0, *, \circ, \bullet)$ matrix that has been constructed in Theorem 3.1. Each A_i has the form $H_n(\lambda)$, or $\mu\Delta_n$, or $J_n(0)$, and so there are 9 types of diagonal blocks $\mathcal{D}(A_i) = \mathcal{D}_{ii}$ and pairs of off-diagonal blocks $\mathcal{D}(A_i, A_j) = (\mathcal{D}_{ji}, \mathcal{D}_{ij})$, $i < j$; they have been given in Theorem 2.2. It suffices to prove that (10)–(17) satisfy the conditions (i) and (ii) from Lemma 4.3.

5 Diagonal blocks of \mathcal{D}

Fist we verify that the diagonal blocks of \mathcal{D} defined in part (i) of Theorem 2.2 satisfy the condition (i) of Lemma 4.3.

5.1 Diagonal blocks $\mathcal{D}(H_n(\lambda))$ with $|\lambda| > 1$

Due to Lemma 4.3(i), it suffices to prove that each $2n$ -by- $2n$ matrix $A = [A_{ij}]_{i,j=1}^2$ can be reduced to exactly one matrix of the form (10) by adding

$$\begin{aligned} & \begin{bmatrix} S_{11}^* & S_{21}^* \\ S_{12}^* & S_{22}^* \end{bmatrix} \begin{bmatrix} 0 & I_n \\ J_n(\lambda) & 0 \end{bmatrix} + \begin{bmatrix} 0 & I_n \\ J_n(\lambda) & 0 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \\ & = \begin{bmatrix} S_{21}^* J_n(\lambda) + S_{21} & S_{11}^* + S_{22} \\ S_{22}^* J_n(\lambda) + J_n(\lambda) S_{11} & S_{12}^* + J_n(\lambda) S_{12} \end{bmatrix} \quad (32) \end{aligned}$$

in which $S = [S_{ij}]_{i,j=1}^{2n}$ is an arbitrary $2n$ -by- $2n$ matrix. Taking $S_{22} = -A_{12}$ and the other $S_{ij} = 0$, we obtain a new matrix A with $A_{12} = 0$. To preserve A_{12} , we hereafter must take S with $S_{11}^* + S_{22} = 0$. Therefore, we can add $S_{21}^* J_n(\lambda) + S_{21}$ to (the new) A_{11} , $S_{12}^* + J_n(\lambda) S_{12}$ to A_{22} , and $S_{11} J_n(\lambda) + J_n(\lambda) S_{11}$ to A_{21} . Using these additions, we can reduce A to the form (10) due to the following 3 lemmas.

Lemma 5.1. *Adding $SJ_n(\lambda) + S^*$, in which λ is a fixed complex number, $|\lambda| \neq 1$, and S is arbitrary, we can reduce each n -by- n matrix to the zero matrix.*

Proof. Let $A = [a_{ij}]$ be an arbitrary n -by- n matrix. We will reduce it along its *skew diagonals* starting from the upper left hand corner:

$$a_{11}, (a_{21}, a_{12}), (a_{31}, a_{22}, a_{13}), \dots, a_{nn}, \quad (33)$$

by adding $\Delta A := SJ_n(\lambda) + S^*$ in which $S = [s_{ij}]$ is any n -by- n matrix. For instance, if $n = 4$ then ΔA is

$$\begin{bmatrix} \lambda s_{11} + 0 + \bar{s}_{11} & \lambda s_{12} + s_{11} + \bar{s}_{21} & \lambda s_{13} + s_{12} + \bar{s}_{31} & \lambda s_{14} + s_{13} + \bar{s}_{41} \\ \lambda s_{21} + 0 + \bar{s}_{12} & \lambda s_{22} + s_{21} + \bar{s}_{22} & \lambda s_{23} + s_{22} + \bar{s}_{32} & \lambda s_{24} + s_{23} + \bar{s}_{42} \\ \lambda s_{31} + 0 + \bar{s}_{13} & \lambda s_{32} + s_{31} + \bar{s}_{23} & \lambda s_{33} + s_{32} + \bar{s}_{33} & \lambda s_{34} + s_{33} + \bar{s}_{43} \\ \lambda s_{41} + 0 + \bar{s}_{14} & \lambda s_{42} + s_{41} + \bar{s}_{24} & \lambda s_{43} + s_{42} + \bar{s}_{34} & \lambda s_{44} + s_{43} + \bar{s}_{44} \end{bmatrix}.$$

We reduce A to 0 by induction: Assume that the first $t-1$ skew diagonals of A are zero. To preserve them, we take the first $t-1$ skew diagonals of S equalling zero. If the t^{th} skew diagonal of S is (x_1, \dots, x_r) , then we can add

$$(\lambda x_1 + \bar{x}_r, \lambda x_2 + \bar{x}_{r-1}, \lambda x_3 + \bar{x}_{r-2}, \dots, \lambda x_r + \bar{x}_1) \quad (34)$$

to the t^{th} skew diagonal of A . Let us show that each vector $(c_1, \dots, c_r) \in \mathbb{C}^r$ is represented in the form (34); that is, the corresponding system of linear equations

$$\lambda x_1 + \bar{x}_r = c_1, \dots, \lambda x_j + \bar{x}_{r-j+1} = c_j, \dots, \lambda x_r + \bar{x}_1 = c_r \quad (35)$$

has a solution. This is clear if $\lambda = 0$. Suppose that $\lambda \neq 0$.

Let $r = 2k + 1$. By (35), $x_j = \lambda^{-1}(c_j - \bar{x}_{r-j+1})$. Replace j by $r - j + 1$:

$$x_{r-j+1} = \lambda^{-1}(c_{r-j+1} - \bar{x}_j). \quad (36)$$

Substituting (36) into the first $k + 1$ equations of (35), we obtain

$$\lambda x_j + \bar{\lambda}^{-1}(\bar{c}_{r-j+1} - x_j) = (\lambda - \bar{\lambda}^{-1})x_j + \bar{\lambda}^{-1}\bar{c}_{r-j+1} = c_j, \quad j = 1, \dots, k + 1.$$

Since $|\lambda| \neq 1$, $\lambda - \bar{\lambda}^{-1} \neq 0$ and we have

$$x_j = \frac{c_j - \bar{\lambda}^{-1}\bar{c}_{r-j+1}}{\lambda - \bar{\lambda}^{-1}} = \frac{\bar{\lambda}c_j - \bar{c}_{r-j+1}}{\lambda\bar{\lambda} - 1}, \quad j = 1, \dots, k + 1. \quad (37)$$

The equalities (36) and (37) give a solution of (35).

If $r = 2k$, then (35) is solved in the same way, but we take $j = 1, \dots, k$ in (37). \square

Lemma 5.2. *Adding $J_n(\lambda)R + R^*$, in which λ is a fixed complex number, $|\lambda| \neq 1$, and R is arbitrary, we can reduce each n -by- n matrix to the zero matrix.*

Proof. By Lemma 5.1, for each n -by- n matrix B there exists S such that $B + SJ_n(\lambda) + S^* = 0$. Then

$$B^* + J_n(\lambda)^* S^* + S = 0.$$

Write

$$Z := \begin{bmatrix} 0 & & 1 \\ & \ddots & \\ 1 & & 0 \end{bmatrix}.$$

Because $ZJ_n(\lambda)^*Z = J_n(\bar{\lambda})$, we have

$$ZB^*Z + J_n(\bar{\lambda})(ZSZ)^* + ZSZ = 0.$$

This ensures Lemma 5.2 since ZB^*Z is arbitrary. \square

Lemma 5.3. *Adding $SJ_n(\lambda) + J_n(\lambda)S$, we can reduce each $n \times n$ matrix to exactly one matrix of the form 0^\vee .*

Proof. Let $A = [a_{ij}]$ be an arbitrary $n \times n$ matrix. Adding

$$\begin{aligned} SJ_n(\lambda) - J_n(\lambda)S &= SJ_n(0) - J_n(0)S \\ &= \begin{bmatrix} s_{21} - 0 & s_{22} - s_{11} & s_{23} - s_{12} & \dots & s_{2n} - s_{1,n-1} \\ \dots & \dots & \dots & \dots & \dots \\ s_{n-1,1} - 0 & s_{n-1,2} - s_{n-2,1} & s_{n-1,3} - s_{n-2,2} & \dots & s_{n-1,n} - s_{n-2,n-1} \\ s_{n1} - 0 & s_{n2} - s_{n-1,1} & s_{n3} - s_{n-1,2} & \dots & s_{nn} - s_{n-1,n-1} \\ 0 - 0 & 0 - s_{n1} & 0 - s_{n2} & \dots & 0 - s_{n,n-1} \end{bmatrix}, \end{aligned}$$

we reduce A along the diagonals

$$a_{n1}, (a_{n-1,1}, a_{n2}), (a_{n-2,1}, a_{n-1,2}, a_{n3}), \dots, a_{1n}$$

to the form 0^\swarrow . □

5.2 Diagonal blocks $\mathcal{D}(\mu\Delta_n)$ with $|\mu| = 1$

Due to Lemma 4.3(i), it suffices to prove that each $n \times n$ matrix A can be reduced to exactly one matrix of the form 0^\searrow if $\mu \notin \mathbb{R}$ or 0^\swarrow if $\mu \in i\mathbb{R}$ by adding $\Delta A := \mu(S^* \Delta_n + \Delta_n S)$ in which $S = [s_{ij}]$ is any n -by- n matrix.

For example, if $n = 4$ then ΔA is

$$\mu \begin{bmatrix} \bar{s}_{41} + s_{41} + i(0 + 0) & \bar{s}_{31} + s_{42} + i(\bar{s}_{41} + 0) & \dots & \bar{s}_{11} + s_{44} + i(\bar{s}_{21} + 0) \\ \bar{s}_{42} + s_{31} + i(0 + s_{41}) & \bar{s}_{32} + s_{32} + i(\bar{s}_{42} + s_{42}) & \dots & \bar{s}_{12} + s_{34} + i(\bar{s}_{22} + s_{44}) \\ \bar{s}_{43} + s_{21} + i(0 + s_{31}) & \bar{s}_{33} + s_{22} + i(\bar{s}_{43} + s_{32}) & \dots & \bar{s}_{13} + s_{24} + i(\bar{s}_{23} + s_{34}) \\ \bar{s}_{44} + s_{11} + i(0 + s_{21}) & \bar{s}_{34} + s_{12} + i(\bar{s}_{44} + s_{22}) & \dots & \bar{s}_{14} + s_{14} + i(\bar{s}_{24} + s_{24}) \end{bmatrix}.$$

Let $\Delta A = \mu[\delta_{ij}]$. Write

$$s_{n+1,j} := 0 \quad \text{for } j = 1, \dots, n. \quad (38)$$

Then

$$\delta_{ij} = \bar{s}_{n+1-j,i} + s_{n+1-i,j} + i(\bar{s}_{n+2-j,i} + s_{n+2-i,j}). \quad (39)$$

Step 1: Let us prove that

$$\exists S: \quad A + \Delta A \text{ is a diagonal matrix.} \quad (40)$$

Let $A = \mu[a_{ij}]$. We need to prove that the system of equations

$$\delta_{ij} = -a_{ij}, \quad i, j = 1, \dots, n, \quad i \neq j \quad (41)$$

with unknowns s_{ij} is consistent for all a_{ij} .

Since

$$\bar{\delta}_{ji} = \bar{s}_{n+2-j,i} + s_{n+2-i,j} - i(\bar{s}_{n+2-j,i} + s_{n+2-i,j}) = -\bar{a}_{ji}$$

we have

$$\begin{aligned} \bar{s}_{n+1-j,i} + s_{n+1-i,j} &= (\delta_{ij} + \bar{\delta}_{ji})/2 = (-a_{ij} - \bar{a}_{ji})/2 \\ \bar{s}_{n+2-j,i} + s_{n+2-i,j} &= (\delta_{ij} - \bar{\delta}_{ji})/(2i) = (-a_{ij} + \bar{a}_{ji})/(2i) \end{aligned}$$

Thus, the system of equations (41) is equivalent to the system

$$\begin{aligned} \bar{s}_{n+1-j,i} + s_{n+1-i,j} &= b_{ij} \\ \bar{s}_{n+2-j,i} + s_{n+2-i,j} &= c_{ij} \end{aligned} \quad i, j = 1, \dots, n, \quad i < j. \quad (42)$$

in which

$$b_{ij} := (-a_{ij} - \bar{a}_{ji})/2, \quad c_{ij} := (-a_{ij} + \bar{a}_{ji})/(2i).$$

For $k, l = 1, \dots, n$, write

$$u_{kl} := \begin{cases} -s_{kl} & \text{if } k + l \geq n + 2, \\ \bar{s}_{kl} & \text{if } k + l \leq n + 1. \end{cases} \quad (43)$$

Then the system (42) takes the form

$$\begin{aligned} u_{n+1-j,i} - u_{n+1-i,j} &= b_{ij} \\ u_{n+2-j,i} - u_{n+2-i,j} &= c_{ij} \end{aligned} \quad i, j = 1, \dots, n, \quad i < j. \quad (44)$$

Rewrite it as follows

$$\begin{aligned} u_{kl} - u_{pq} &= b'_{kl}, & k + q = l + p = n + 1, & \quad k < p, \\ u_{kl} - u_{p'q'} &= c'_{kl}, & k + q' = l + p' = n + 2, & \quad k < p'. \end{aligned} \quad (45)$$

Since $k - l = p - q = p' - q'$, the system (45) is partitioned into subsystems with unknowns u_{ij} , $i - j = \text{const}$. Each of these subsystems has the form

$$\dots, u_{kl} - u_{p+1,q+1} = c'_{kl}, \quad u_{kl} - u_{pq} = b'_{kl}, \quad u_{k+1,l+1} - u_{pq} = c'_{k+1,l+1}, \quad \dots \quad (46)$$

and is consistent. This proves (40).

Step 2: Let us prove that for each diagonal matrix A

$$\exists S : \quad A + \Delta A \text{ has the form } 0^{\setminus} \text{ if } \mu \notin \mathbb{R} \text{ or } 0^{\setminus} \text{ if } \mu \notin i\mathbb{R}. \quad (47)$$

Since A , 0^{\setminus} , and 0^{\setminus} are diagonal, the matrix ΔA must be diagonal too. Thus, the entries of S must satisfy the system (41) with $a_{ij} = 0$. Reasoning as in Step 1, we obtain the system (45) with $b'_{kl} = c'_{kl} = 0$, which is partitioned into subsystems (46). Each of these subsystems is represented in the form

$$u_{1,r+1} = u_{2,r+2} = \dots = u_{n-r,n} \quad (48)$$

in which $r \geq 0$, or

$$u_{r+1,1} = u_{r+2,2} = \dots = u_{n,n-r} = u_{n+1,n-r+1} = 0 \quad (\text{see (38)}) \quad (49)$$

in which $r \geq 1$. By (43), S is upper triangular and

$$s_{1,r+1} = \cdots = s_{z,r+z} = -\bar{s}_{z+1,r+z+1} = \cdots = -\bar{s}_{n-r,n}$$

in which z is the integer part of $(n+1-r)/2$ and $r = 0, 1, \dots, n-2$.

Let $n = 2m$ or $2m+1$. By (39), the first m entries of the main diagonal of $\mu^{-1}\Delta A$ are

$$\begin{aligned} & \bar{s}_{n1} + s_{n1} \\ & s_{n-1,2} + s_{n-1,2} + i(\bar{s}_{n2} + s_{n2}) \\ & \dots\dots\dots \\ & \bar{s}_{n+1-m,m} + s_{n+1-m,m} + i(\bar{s}_{n+2-m,m} + s_{n+2-m,m}). \end{aligned}$$

They are zero and so we cannot change the first m diagonal entries of A .

The last m entries of the main diagonal of $\mu^{-1}\Delta A$ are

$$\begin{aligned} & \bar{s}_{m,n-m+1} + s_{m,n-m+1} + i(\bar{s}_{m+1,n-m+1} + s_{m+1,n-m+1}) \\ & \dots\dots\dots \\ & \bar{s}_{2,n-1} + s_{2,n-1} + i(\bar{s}_{3,n-1} + s_{3,n-1}) \\ & \bar{s}_{1n} + s_{1n} + i(\bar{s}_{2n} + s_{2n}). \end{aligned}$$

They are arbitrary and we make zero the last m entries of the main diagonal of A . This proves (47) for $n = 2m$.

Let $n = 2m+1$. Since $s_{m+2,m+1} = 0$, the $(m+1)$ st entry of $\mu^{-1}\Delta A$ is

$$\delta_{m+1,m+1} = \bar{s}_{m+1,m+1} + s_{m+1,m+1},$$

which is an arbitrary real number. Thus, we can add μr with an arbitrary $r \in \mathbb{R}$ to the $(m+1)$ st entry of A . This proves (47) for $n = 2m+1$.

5.3 Diagonal blocks $\mathcal{D}(J_n(0))$

Due to Lemma 4.3(i), it suffices to prove that each n -by- n matrix A can be reduced to exactly one matrix of the form 0^\wedge by adding

$$\begin{aligned} \Delta A &:= S^* J_n(0) + J_n(0) S \\ &= \begin{bmatrix} 0 + s_{21} & \bar{s}_{11} + s_{22} & \bar{s}_{21} + s_{23} & \dots & \bar{s}_{n-1,1} + s_{2n} \\ 0 + s_{31} & \bar{s}_{12} + s_{32} & \bar{s}_{22} + s_{33} & \dots & \bar{s}_{n-1,2} + s_{3n} \\ \dots\dots\dots \\ 0 + s_{n1} & \bar{s}_{1,n-1} + s_{n2} & \bar{s}_{2,n-1} + s_{n3} & \dots & \bar{s}_{n-1,n-1} + s_{nn} \\ 0 + 0 & \bar{s}_{1n} + 0 & \bar{s}_{2n} + 0 & \dots & \bar{s}_{n-1,n} + 0 \end{bmatrix}, \quad (50) \end{aligned}$$

in which $S = [s_{ij}]$ is any n -by- n matrix. Since

$$\Delta A = [b_{ij}], \quad b_{ij} := \bar{s}_{j-1,i} + s_{i+1,j} \quad (s_{0i} := 0, \quad s_{n+1,j} := 0), \quad (51)$$

all entries of ΔA have the form $\bar{s}_{kl} + s_{l+1,k+1}$. The transitive closure of $(k, l) \sim (l+1, k+1)$ is an equivalence relation on the set $\{1, \dots, n\} \times \{1, \dots, n\}$. Decompose ΔA into the sum of matrices

$$\Delta A = B_{n1} + B_{n-1,1} + \dots + B_{11} + B_{12} + \dots + B_{1,n-1}$$

that correspond to the equivalence classes and are defined as follows. Each B_{1j} ($j = 1, 2, \dots, n$) is obtained from ΔA by replacing with 0 all of its entries except for

$$\bar{s}_{1j} + s_{j+1,2}, \quad \bar{s}_{j+1,2} + s_{3,j+2}, \quad \bar{s}_{3,j+2} + s_{j+3,4}, \quad \dots \quad (52)$$

and each B_{i1} ($i = 2, 3, \dots, n$) is obtained from ΔA by replacing with 0 all of its entries except for

$$0 + s_{i1}, \quad \bar{s}_{i1} + s_{2,i+1}, \quad \bar{s}_{2,i+1} + s_{i+2,3}, \quad \bar{s}_{i+2,3} + s_{4,i+3}, \quad \bar{s}_{4,i+3} + s_{i+4,5}, \quad \dots; \quad (53)$$

the pairs of indices in (52) and in (53) are equivalent:

$$(1, j) \sim (j+1, 2) \sim (3, j+2) \sim (j+3, 4) \sim \dots$$

and

$$(i, 1) \sim (2, i+1) \sim (i+2, 3) \sim (4, i+3) \sim (i+4, 5) \sim \dots$$

We call the entries (52) and (53) the *main entries* of B_{1j} and B_{i1} ($i > 1$). The matrices $B_{n1}, \dots, B_{11}, B_{12}, \dots, B_{1n}$ have no common s_{ij} , and so we can add to A each of these matrices separately.

The entries of the sequence (52) are independent: an arbitrary sequence of complex numbers can be represented in the form (52). The entries (53) are dependent only if the last entry in this sequence has the form $\bar{s}_{kn} + 0$ (see (50)); then $(k, n) = (2p, i-1+2p)$ for some p , and so $i = n+1-2p$. Thus the following sequences (53) are dependent:

$$\begin{aligned} & 0 + s_{n-1,1}, \quad \bar{s}_{n-1,1} + s_{2n}, \quad \bar{s}_{2n} + 0; \\ & 0 + s_{n-3,1}, \quad \bar{s}_{n-3,1} + s_{2,n-2}, \quad \bar{s}_{2,n-2} + s_{n-1,3}, \quad \bar{s}_{n-1,3} + s_{4n}, \quad \bar{s}_{4n} + 0; \quad \dots \end{aligned}$$

One of the main entries of each of the matrices $B_{n-1,1}, B_{n-3,1}, B_{n-5,1}, \dots$ is expressed through the other main entries of this matrix, which are arbitrary. The main entries of the other matrices B_{i1} and B_{1j} are arbitrary. Adding B_{i1} and B_{1j} , we reduce A to the form 0^\wedge .

6 Off-diagonal blocks of \mathcal{D} that correspond to summands of A_{can} of the same type

Now we verify the condition (ii) of Lemma 4.3 for those off-diagonal blocks of \mathcal{D} (defined in Theorem 2.2(ii)) whose horizontal and vertical strips contain summands of A_{can} of the same type.

6.1 Pairs of blocks $\mathcal{D}(H_m(\lambda), H_n(\mu))$ with $|\lambda|, |\mu| > 1$

Due to Lemma 4.3(ii), it suffices to prove that each pair (B, A) of $2n \times 2m$ and $2m \times 2n$ matrices can be reduced to exactly one pair of the form (13) by adding

$$(S^* H_m(\lambda) + H_n(\mu) R, R^* H_n(\mu) + H_m(\lambda) S), \quad S \in \mathbb{C}^{m \times n}, \quad R \in \mathbb{C}^{n \times m}.$$

Putting $R = 0$ and $S = -H_m(\lambda)^{-1} A$, we reduce A to 0. To preserve $A = 0$ we hereafter must take S and R such that $R^* H_n(\mu) + H_m(\lambda) S = 0$; that is,

$$S = -H_m(\lambda)^{-1} R^* H_n(\mu),$$

and so we can add

$$\Delta B := -H_n(\mu)^* R H_m(\lambda)^{-*} H_m(\lambda) + H_n(\mu) R$$

to B .

Write $P := -H_n(\mu)^* R$, then $R = -H_n(\mu)^{-*} P$ and

$$\Delta B = P \begin{bmatrix} J_m(\lambda) & 0 \\ 0 & J_m(\bar{\lambda})^{-T} \end{bmatrix} - \begin{bmatrix} J_n(\bar{\mu})^{-T} & 0 \\ 0 & J_n(\mu) \end{bmatrix} P \quad (54)$$

Partition B , ΔB , and P into $n \times m$ blocks:

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \quad \Delta B = \begin{bmatrix} \Delta B_{11} & \Delta B_{12} \\ \Delta B_{21} & \Delta B_{22} \end{bmatrix}, \quad P = \begin{bmatrix} X & Y \\ Z & T \end{bmatrix}.$$

By (54),

$$\begin{aligned} \Delta B_{11} &= X J_m(\lambda) - J_n(\bar{\mu})^{-T} X, & \Delta B_{12} &= Y J_m(\bar{\lambda})^{-T} - J_n(\bar{\mu})^{-T} Y, \\ \Delta B_{21} &= Z J_m(\lambda) - J_n(\mu) Z, & \Delta B_{22} &= T J_m(\bar{\lambda})^{-T} - J_n(\mu) T. \end{aligned}$$

These equalities ensure that we can reduce each block B_{ij} separately by adding ΔB_{ij} .

(i) First we reduce B_{11} by adding $\Delta B_{11} = X J_m(\lambda) - J_n(\bar{\mu})^{-T} X$.

Since $|\lambda| > 1$ and $|\mu| > 1$, we have that $J_m(\lambda)$ and $J_n(\bar{\mu})^{-T}$ have no common eigenvalues and so ΔB_{11} is an arbitrary matrix. We make $B_{11} = 0$.

(ii) Let us reduce B_{12} by adding $\Delta B_{12} = Y J_m(\bar{\lambda})^{-T} - J_n(\bar{\mu})^{-T} Y$.

If $\lambda \neq \mu$, then ΔB_{12} is arbitrary; we make $B_{12} = 0$.

Let $\lambda = \mu$. Write $F := J_n(0)$. Since

$$J_n(\bar{\lambda})^{-1} = (\bar{\lambda} I_n + F)^{-1} = \bar{\lambda}^{-1} I_n - \bar{\lambda}^{-2} F + \bar{\lambda}^{-3} F^2 - \dots,$$

we have

$$\begin{aligned} \Delta B_{12} &= Y (J_m(\bar{\lambda})^{-T} - \bar{\lambda}^{-1} I_m) - (J_n(\bar{\lambda})^{-T} - \bar{\lambda}^{-1} I_n) Y \\ &= -\bar{\lambda}^{-2} \begin{bmatrix} y_{12} & \dots & y_{1m} & 0 \\ y_{22} & \dots & y_{2m} & 0 \\ y_{32} & \dots & y_{3m} & 0 \\ \dots & \dots & \dots & \dots \end{bmatrix} + \bar{\lambda}^{-2} \begin{bmatrix} 0 & \dots & 0 \\ y_{11} & \dots & y_{1m} \\ y_{21} & \dots & y_{2m} \\ \dots & \dots & \dots \end{bmatrix} + \dots \end{aligned}$$

We reduce B_{12} to the form 0^{\nearrow} along its diagonals starting from the upper right hand corner.

(iii) Let us reduce B_{21} by adding $\Delta B_{21} = Z J_m(\lambda) - J_n(\mu) Z$.

If $\lambda \neq \mu$, then ΔB_{21} is arbitrary; we make $B_{21} = 0$.

If $\lambda = \mu$, then

$$\begin{aligned} \Delta B_{21} &= Z (J_m(\lambda) - \lambda I_m) - (J_n(\lambda) - \lambda I_n) Z \\ &= \begin{bmatrix} 0 & z_{11} & \dots & z_{1,m-1} \\ \dots & \dots & \dots & \dots \\ 0 & z_{n-1,1} & \dots & z_{n-1,m-1} \\ 0 & z_{n1} & \dots & z_{n,m-1} \end{bmatrix} - \begin{bmatrix} z_{21} & \dots & z_{2m} \\ \dots & \dots & \dots \\ z_{n1} & \dots & z_{nm} \\ 0 & \dots & 0 \end{bmatrix}; \end{aligned}$$

we reduce B_{12} to the form 0^{\swarrow} along its diagonals starting from the lower left hand corner.

(iv) Finally, reduce B_{22} by adding $\Delta B_{22} = T J_m(\bar{\lambda})^{-T} - J_n(\mu) T$.— Since $|\lambda| > 1$ and $|\mu| > 1$, ΔB_{22} is arbitrary; we make $B_{22} = 0$.

6.2 Pairs of blocks $\mathcal{D}(\mu\Delta_m, \nu\Delta_n)$ with $|\mu| = |\nu| = 1$

Due to Lemma 4.3(ii), it suffices to prove that each pair (B, A) of $n \times m$ and $m \times n$ matrices can be reduced to exactly one pair of the form $(0, 0)$ if $\mu \neq \pm\nu$ and $(0^\times, 0)$ if $\mu = \pm\nu$ by adding

$$(\mu S^* \Delta_m + \nu \Delta_n R, \nu R^* \Delta_n + \mu \Delta_m S), \quad S \in \mathbb{C}^{m \times n}, \quad R \in \mathbb{C}^{n \times m}.$$

We put $R = 0$ and $S = -\bar{\mu}\Delta_m^{-1}A$, which reduces A to 0. To preserve $A = 0$ we hereafter must take S and R such that $\nu R^* \Delta_n + \mu \Delta_m S = 0$; that is, $S = -\bar{\mu}\nu\Delta_m^{-1}R^* \Delta_n$, and so we can add

$$\Delta B := \nu \Delta_n R - \mu^2 \bar{\nu} \Delta_n^* R \Delta_m^{-*} \Delta_m$$

to B .

Write $P := \Delta_n^* R$, then

$$\Delta B = \bar{\nu}[\nu^2(\Delta_n \Delta_n^{-*})P - \mu^2 P(\Delta_m^{-*} \Delta_m)]. \quad (55)$$

Since

$$\Delta_n^{-*} = \begin{bmatrix} * & & i & 1 \\ & \ddots & \ddots & \\ i & 1 & & \\ 1 & & & 0 \end{bmatrix},$$

we have

$$\Delta_n \Delta_n^{-*} = \begin{bmatrix} 1 & & & 0 \\ 2i & 1 & & \\ & \ddots & \ddots & \\ * & & 2i & 1 \end{bmatrix} \quad (56)$$

and

$$\Delta_m^{-*} \Delta_m = (\Delta_n \Delta_n^{-*})^T = \begin{bmatrix} 1 & 2i & & * \\ & 1 & \ddots & \\ & & \ddots & 2i \\ 0 & & & 1 \end{bmatrix}. \quad (57)$$

If $\mu \neq \pm\nu$, then $\mu^2 \neq \nu^2$, the matrices $\nu^2(\Delta_n \Delta_n^{-*})$ and $\mu^2(\Delta_m^{-*} \Delta_m)$ have distinct eigenvalues, and so ΔB can be made arbitrary. We make $B = 0$.

If $\mu = \pm\nu$, then

$$\frac{1}{2i\nu} \Delta B = \begin{bmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ * & & 1 & 0 \end{bmatrix} P - P \begin{bmatrix} 0 & 1 & & * \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{bmatrix},$$

and we reduce B to the form 0^\lessgtr along its skew diagonals starting from the upper left hand corner.

6.3 Pairs of blocks $\mathcal{D}(J_m(0), J_n(0))$ with $m \geq n$.

Due to Lemma 4.3(ii), it suffices to prove that each pair (B, A) of $n \times m$ and $m \times n$ matrices with $m \geq n$ can be reduced to exactly one pair of the form $(0^\lessgtr, 0^\lessgtr)$ if n is even and of the form $(0^\lessgtr + \mathcal{P}_{nm}, 0^\lessgtr)$ if n is odd by adding the matrices

$$\Delta A = R^* J_n(0) + J_m(0) S, \quad \Delta B^* = J_m(0)^T S + R^* J_n(0)^T \quad (58)$$

to A and B^* (we prefer to reduce B^* instead of B).

Write $S = [s_{ij}]$ and $R^* = [-r_{ij}]$ (they are m -by- n). Then

$$\Delta A = \begin{bmatrix} s_{21} - 0 & s_{22} - r_{11} & s_{23} - r_{12} & \dots & s_{2n} - r_{1,n-1} \\ \dots & \dots & \dots & \dots & \dots \\ s_{m-1,1} - 0 & s_{m-1,2} - r_{m-2,1} & s_{m-1,3} - r_{m-2,2} & \dots & s_{m-1,n} - r_{m-2,n-1} \\ s_{m1} - 0 & s_{m2} - r_{m-1,1} & s_{m3} - r_{m-1,2} & \dots & s_{mn} - r_{m-1,n-1} \\ 0 - 0 & 0 - r_{m1} & 0 - r_{m2} & \dots & 0 - r_{m,n-1} \end{bmatrix}$$

and

$$\Delta B^* = \begin{bmatrix} 0 - r_{12} & 0 - r_{13} & \dots & 0 - r_{1n} & 0 - 0 \\ s_{11} - r_{22} & s_{12} - r_{23} & \dots & s_{1,n-1} - r_{2n} & s_{1n} - 0 \\ \dots & \dots & \dots & \dots & \dots \\ s_{m-2,1} - r_{m-1,2} & s_{m-2,2} - r_{m-1,3} & \dots & s_{m-2,n-1} - r_{m-1,n} & s_{m-2,n} - 0 \\ s_{m-1,1} - r_{m2} & s_{m-1,2} - r_{m3} & \dots & s_{m-1,n-1} - r_{mn} & s_{m-1,n} - 0 \end{bmatrix}.$$

Adding ΔA , we reduce A to the form

$$0^\lessgtr := \begin{bmatrix} 0_{m-1,n} \\ * & * & \dots & * \end{bmatrix}. \quad (59)$$

To preserve this form, we hereafter must take

$$s_{21} = \dots = s_{m1} = 0, \quad s_{ij} = r_{i-1,j-1} \quad (2 \leq i \leq m, 2 \leq j \leq n). \quad (60)$$

Write

$$(r_{00}, r_{01}, \dots, r_{0,n-1}) := (s_{11}, s_{12}, \dots, s_{1n}),$$

then

$$\Delta B^* = \begin{bmatrix} 0 - r_{12} & 0 - r_{13} & \dots & 0 - r_{1n} & 0 - 0 \\ r_{00} - r_{22} & r_{01} - r_{23} & \dots & r_{0,n-2} - r_{2n} & r_{0,n-1} - 0 \\ 0 - r_{32} & r_{11} - r_{33} & \dots & r_{1,n-2} - r_{3n} & r_{1,n-1} - 0 \\ 0 - r_{42} & r_{21} - r_{43} & \dots & r_{2,n-2} - r_{4n} & r_{2,n-1} - 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 - r_{m2} & r_{m-2,1} - r_{m3} & \dots & r_{m-2,n-2} - r_{mn} & r_{m-2,n-1} - 0 \end{bmatrix}. \quad (61)$$

If r_{ij} and $r_{i'j'}$ are parameters of the same diagonal of ΔB^* , then $i - j = i' - j'$. Hence, the diagonals of ΔB^* have no common parameters, and so we can reduce the diagonals of B^* independently.

The first n diagonals of ΔB^* starting from the upper right hand corner are

$$\begin{aligned} &0, \quad (-r_{1n}, r_{0,n-1}), \quad (\underline{-r_{1,n-1}}, r_{0,n-2} - r_{2n}, \underline{r_{1,n-1}}), \\ &\quad (-r_{1,n-2}, r_{0,n-3} - r_{2,n-1}, r_{1,n-2} - r_{3n}, r_{2,n-1}), \\ &(\underline{-r_{1,n-3}}, r_{0,n-4} - r_{2,n-2}, \underline{r_{1,n-3} - r_{3,n-1}}, r_{2,n-2} - r_{4n}, \underline{r_{3,n-1}}), \dots \end{aligned}$$

(we underline linearly dependent entries), adding them we reduce the first n diagonals of B^* to the form $0^<$.

The $(n+1)^{\text{st}}$ diagonal of ΔB^* is

$$\begin{cases} (r_{00} - r_{22}, r_{11} - r_{33}, \dots, r_{n-2,n-2} - r_{nn}) & \text{if } m = n, \\ (r_{00} - r_{22}, r_{11} - r_{33}, \dots, r_{n-2,n-2} - r_{nn}, r_{n-1,n-1}) & \text{if } m > n. \end{cases}$$

Adding it, we make the $(n+1)^{\text{st}}$ diagonal of B^* zero.

If $m > n+1$, then the $(n+2)^{\text{nd}}, \dots, m^{\text{th}}$ diagonals of ΔB^* are

$$\begin{aligned} &(\underline{-r_{32}}, r_{21} - r_{43}, \underline{r_{32} - r_{54}}, \dots, r_{n,n-1}), \\ &\dots\dots\dots \\ &(\underline{-r_{m-n+1,2}}, r_{m-n,1} - r_{m-n+2,3}, \underline{r_{m-n+1,2} - r_{m-n+3,4}}, \dots, r_{m-2,n-1}). \end{aligned}$$

Each of these diagonals contains n elements. If n is even, then the length of each diagonal is even and its elements are linearly independent; we make the corresponding diagonals of B^* equal to zero. If n is odd, then the length of each diagonal is odd and the set of its odd-numbered elements is linearly dependent; we make all elements of the corresponding diagonals of B^* equal

to zero except for their last elements (they correspond to the stars of \mathcal{P}_{nm} , which is defined in (6)).

It remains to reduce the last $n-1$ diagonals of B^* (the last $n-2$ diagonals if $m = n$). The corresponding diagonals of ΔB^* are

$$\begin{aligned} & -r_{m2}, \\ & (-r_{m-1,2}, r_{m-2,1} - r_{m3}), \\ & (-r_{m-2,2}, r_{m-3,1} - r_{m-1,3}, r_{m-2,2} - r_{m4}), \\ & (-r_{m-3,2}, r_{m-4,1} - r_{m-2,3}, r_{m-3,2} - r_{m-1,4}, r_{m-2,3} - r_{m5}), \\ & \dots\dots\dots \\ & (-r_{m-n+3,2}, r_{m-n+2,1} - r_{m-n+4,3}, \dots, r_{m-2,n-3} - r_{m,n-1}), \end{aligned}$$

and, only if $m > n$,

$$(-r_{m-n+2,2}, r_{m-n+1,1} - r_{m-n+3,3}, \dots, r_{m-2,n-2} - r_{mn}).$$

Adding these diagonals, we make the corresponding diagonals of B^* zero. To preserve the zero diagonals, we hereafter must take $r_{m2} = r_{m4} = r_{m6} = \dots = 0$ and arbitrary $r_{m1}, r_{m3}, r_{m5}, \dots$.

Recall that A has the form 0^\downarrow (see (59)). Since $r_{m1}, r_{m3}, r_{m5}, \dots$ are arbitrary, we can reduce A to the form

$$\begin{bmatrix} 0_{m-1,n} \\ * & 0 & * & 0 & \dots \end{bmatrix}$$

by adding ΔA ; these additions preserve B .

If $m = n$, then we may alternatively reduce A to the form

$$\begin{bmatrix} \dots\dots\dots \\ 0 & 0 & \dots & 0 \\ * & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ * & 0 & \dots & 0 \end{bmatrix}$$

preserving the form 0^\wedge of B .

7 Off-diagonal blocks of \mathcal{D} that correspond to summands of A_{can} of distinct types

Finally, we verify the condition (ii) of Lemma 4.3 for those off-diagonal blocks of \mathcal{D} (defined in Theorem 2.2(iii)) whose horizontal and vertical strips contain summands of A_{can} of different types.

7.1 Pairs of blocks $\mathcal{D}(H_m(\lambda), \mu\Delta_n)$ with $|\lambda| > 1$ and $|\mu| = 1$

Due to Lemma 4.3(ii), it suffices to prove that each pair (B, A) of $n \times 2m$ and $2m \times n$ matrices can be reduced to the pair $(0, 0)$ by adding

$$(S^* H_m(\lambda) + \mu\Delta_n R, R^* \mu\Delta_n + H_m(\lambda)S), \quad S \in \mathbb{C}^{2m \times n}, \quad R \in \mathbb{C}^{n \times 2m}.$$

Reduce A to 0 by putting $R = 0$ and $S = -H_m(\lambda)^{-1}A$. To preserve $A = 0$, we hereafter must take S and R such that $R^* \mu\Delta_n + H_m(\lambda)S = 0$; that is,

$$S = -H_m(\lambda)^{-1}R^* \mu\Delta_n.$$

Hence, we can add

$$\Delta B := \mu\Delta_n R - \bar{\mu}\Delta_n^* R H_m(\lambda)^{-*} H_m(\lambda)$$

to B . Write $P = \bar{\mu}\Delta_n^* R$, then

$$\Delta B = \mu\bar{\mu}^{-1}\Delta_n\Delta_n^{-*}P - P(J_m(\lambda) \oplus J_m(\bar{\lambda})^{-T}).$$

By (56), $\mu\bar{\mu}^{-1}\Delta_n\Delta_n^{-*}$ has the single eigenvalue $\mu\bar{\mu}^{-1}$, which is of modulus 1. Since $|\lambda| > 1$, $\mu\bar{\mu}^{-1}\Delta_n\Delta_n^{-*}$ and $J_m(\lambda) \oplus J_m(\bar{\lambda})^{-T}$ have no common eigenvalues. Thus, ΔB is an arbitrary matrix and we make $B = 0$.

7.2 Pairs of blocks $\mathcal{D}(H_m(\lambda), J_n(0))$ with $|\lambda| > 1$

Due to Lemma 4.3(ii), it suffices to prove that each pair (B, A) of $n \times 2m$ and $2m \times n$ matrices can be reduced to exactly one pair of the form $(0, 0)$ if n is even and to the form $(0^\dagger, 0)$ if n is odd by adding

$$(S^* H_m(\lambda) + J_n(0)R, R^* J_n(0) + H_m(\lambda)S), \quad S \in \mathbb{C}^{2m \times n}, \quad R \in \mathbb{C}^{n \times 2m}.$$

Putting $R = 0$ and $S = -H_m(\lambda)^{-1}A$, we reduce A to 0. To preserve $A = 0$ we hereafter must take S and R such that $R^* J_n(0) + H_m(\lambda)S = 0$; that is,

$$S = -H_m(\lambda)^{-1}R^* J_n(0).$$

Hence we can add

$$\begin{aligned} \Delta B &:= J_n(0)R - J_n(0)^T R H_m(\lambda)^{-*} H_m(\lambda) \\ &= J_n(0)R - J_n(0)^T R (J_m(\lambda) \oplus J_m(\bar{\lambda})^{-T}) \end{aligned}$$

to B .

Divide B and R into two blocks of size $n \times m$:

$$B = [M \ N], \quad R = [U \ V].$$

We can add to M and N the matrices

$$\Delta M := J_n(0)U - J_n(0)^T U J_m(\lambda), \quad \Delta N := J_n(0)V - J_n(0)^T V J_m(\bar{\lambda})^{-T}.$$

We reduce M as follows. Let $(u_1, u_2, \dots, u_n)^T$ be the first column of U . Then we can add to the first column \vec{b}_1 of M the vector

$$\begin{aligned} \Delta \vec{b}_1 &:= (u_2, \dots, u_n, 0)^T - \lambda(0, u_1, \dots, u_{n-1})^T \\ &= \begin{cases} 0 & \text{if } n = 1, \\ (u_2, u_3 - \lambda u_1, u_4 - \lambda u_2, \dots, u_n - \lambda u_{n-2}, -\lambda u_{n-1})^T & \text{if } n > 1. \end{cases} \end{aligned}$$

The elements of this vector are linearly independent if n is even, and they are linearly dependent if n is odd. We reduce \vec{b}_1 to zero if n is even, and to the form $(*, 0, \dots, 0)^T$ or $(0, \dots, 0, *)^T$ if n is odd. Then we successively reduce the other columns transforming M to 0 if n is even, and to the form 0_{nm}^\dagger if n is odd.

We reduce N in the same way starting from the last column.

7.3 Pairs of blocks $\mathcal{D}(\lambda \Delta_m, J_n(0))$ with $|\lambda| = 1$

Due to Lemma 4.3(ii), it suffices to prove that each pair (B, A) of $n \times m$ and $m \times n$ matrices can be reduced to exactly one pair of the form $(0, 0)$ if n is even and to the form $(0^\dagger, 0)$ if n is odd by adding

$$(S^* \lambda \Delta_m + J_n(0)R, R^* J_n(0) + \lambda \Delta_m S), \quad S \in \mathbb{C}^{m \times n}, \ R \in \mathbb{C}^{n \times m}.$$

Putting $R = 0$ and $S = -\bar{\lambda} \Delta_m^{-1} A$, we reduce A to 0. To preserve $A = 0$ we hereafter must take S and R such that $R^* J_n(0) + \lambda \Delta_m S = 0$; that is, $S = -\bar{\lambda} \Delta_m^{-1} R^* J_n(0)$. By (57), we can add

$$\begin{aligned} \Delta B &:= J_n(0)R - \lambda^2 J_n(0)^T R \Delta_m^{-*} \Delta_m \\ &= \begin{bmatrix} r_{21} & \dots & r_{2m} \\ \dots & \dots & \dots \\ r_{n1} & \dots & r_{nm} \\ 0 & \dots & 0 \end{bmatrix} - \lambda^2 \begin{bmatrix} 0 & \dots & 0 \\ r_{11} & \dots & r_{1m} \\ \dots & \dots & \dots \\ r_{n-1,1} & \dots & r_{n-1,m} \end{bmatrix} \begin{bmatrix} 1 & 2i & * \\ & 1 & \ddots \\ & & \ddots & 2i \\ 0 & & & 1 \end{bmatrix} \end{aligned}$$

to B . We reduce B to 0 if n is even and to 0^\dagger if n is odd along its columns starting from the first column.

References

- [1] V.I. Arnold, On matrices depending on parameters, *Russian Math. Surveys* 26 (2) (1971) 29–43.
- [2] V. Arnold, Lectures on bifurcations and versal deformations, *Russian Math Surveys* 27 (5) (1972) 54–123.
- [3] V.I. Arnold, *Geometrical methods in the theory of ordinary differential equations*. Springer-Verlag, New York, 1988.
- [4] R. Benedetti, P. Cragolini, Versal families of matrices with respect to unitary conjugation, *Advances in Math.* 54 (1984) 314–335.
- [5] F. De Terán, F.M. Dopico, The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits, *Linear Algebra Appl.* 434 (2011) 44–67.
- [6] F. De Terán, F.M. Dopico, The equation $XA + AX^* = 0$ and the dimension of $*$ -congruence orbits, *Electr. J. Linear Algebra*, 2011 (to appear).
- [7] D.Z. Djokovic, J. Patera, P. Winternitz, H. Zassenhaus, Normal forms of elements of classical real and complex Lie and Jordan algebras, *J. Math. Phys.* 24 (6) (1983) 1363–1374.
- [8] A. Edelman, E. Elmroth, B. Kågström, A geometric approach to perturbation theory of matrices and matrix pencils. Part I: Versal deformations, *Siam J. Matrix Anal. Appl.* 18 (3) (1997) 653–692.
- [9] V. Futorny, V.V. Sergeichuk, Miniversal deformations of matrices of bilinear forms, Preprint RT-MAT 2007-04, Universidade de São Paulo, 2007, 34 p. (arXiv:1004.3584v1).
- [10] V. Futorny, V.V. Sergeichuk, Change of the congruence canonical form of 2×2 and 3×3 matrices under perturbations, Preprint RT-MAT 2007-02, Universidade de São Paulo, 2007, 18 p. (arXiv:1004.3590v1).
- [11] D.M. Galin, On real matrices depending on parameters, *Uspehi Mat. Nauk* 27 (1) (1972) 241–242.
- [12] D.M. Galin, Versal deformations of linear Hamiltonian systems, *Trudy Semin. I. G. Petrovsky* 1 (1975) 63–73 (in Russian).

- [13] M.I. Garcia-Planas, V.V. Sergeichuk, Simplest miniversal deformations of matrices, matrix pencils, and contragredient matrix pencils, *Linear Algebra Appl.* 302–303 (1999) 45–61.
- [14] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge U. P., Cambridge, 1985.
- [15] R.A. Horn, V.V. Sergeichuk, Congruence of a square matrix and its transpose, *Linear Algebra Appl.* 389 (2004) 347–353.
- [16] R.A. Horn, V.V. Sergeichuk, A regularization algorithm for matrices of bilinear and sesquilinear forms, *Linear Algebra Appl.* 412 (2006) 380–395.
- [17] R.A. Horn, V.V. Sergeichuk, Canonical forms for complex matrix congruence and $*$ -congruence, *Linear Algebra Appl.* 416 (2006) 1010–1032.
- [18] R.A. Horn, V.V. Sergeichuk, Canonical matrices of bilinear and sesquilinear forms, *Linear Algebra Appl.* 428 (2008) 193–223.
- [19] H. Kraft, *Geometrische Methoden in der Invariantentheorie*, Friedr. Vieweg & Sohn, Braunschweig, 1985.
- [20] A.I. Markushevich, *Theory of Functions of a Complex Variable*. Vol. I, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1965.
- [21] J. Patera, C. Rousseau, Complex orthogonal and symmetric matrices depending on parameters, *J. Math. Phys.* 23 (5) (1983) 705–714.
- [22] J. Patera, C. Rousseau, D. Schlomiuk, Versal deformations of elements of real classical Lie algebras, *J. Phys. A: Math. Gen* 15 (1982) 1063–1086. (pochitat' kak soslat'sja)
- [23] J. Patera, C. Rousseau, Versal deformations of elements of classical Jordan algebras, *J. Math. Phys.* 24 (6) (1983) 1375–1380.
- [24] V.V. Sergeichuk, Classification problems for system of forms and linear mappings, *Math. USSR, Izvestiya* 31 (3) (1988) 481–501.